



















ods, the linkless analysis method, and the keyword analysis method.

SPOONS plays a valuable and complementary role within its suite of monitoring systems because it does not suffer from the 3 key weaknesses of the existing monitoring systems. It does not share infrastructure with the systems it is monitoring, it monitors 100% of customer interactions, and it does not suffer from the “latency” of human phone calls as an alert mechanism.

## 8.2 Future Work

Even though the current version of some of the SPOONS methods have already been deployed at Netflix, additional challenges remain for future development:

### *Keyword Analysis Improvements.*

Matsuo et al.[8] created a procedure for algorithmically determining the keywords in a document. This procedure could be run on a “document” consisting of tweets that were posted during outage times and updated every time Netflix reported official outage events in the system. This would allow for a larger number of keywords that would dynamically update to new ways of reporting outages over time.

### *Classification Improvements.*

Retraining the classifiers with tweets that occurred during a time identified as an outage would allow us to produce a dynamic training set that may be able to keep up with social trends.

### *Sentiment Analysis.*

The development of a sentiment analysis method is already in progress. This method will determine the sentiment of a tweet and then detect outages by looking for significant increases in negative sentiment.

### *The Nature of an Outage.*

Netflix would like SPOONS to include information in the alert email about the nature of an outage, e.g. which hardware platform is experiencing streaming issues.

### *Malicious Tweet Attacks.*

Currently it is possible for a malicious Twitter user to send a large quantity of tweets “reporting and outage” and trigger false positives in the system. The only existing defense against this kind of attack is that Netflix isn’t going to announce this monitoring system publicly. However, this could possibly be further avoided through the use of author profiling.

## 9. REFERENCES

- [1] A. Culotta. Detecting influenza outbreaks by analyzing twitter messages. In *KDD Workshop on Social Media Analytics*, 2010.
- [2] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, 2009.
- [3] L. Hong and B. D. Davison. Empirical study of topic modeling in twitter. In *Proceedings of the First Workshop on Social Media Analytics, SOMA '10*, pages 80–88, New York, NY, USA, 2010. ACM.
- [4] F. Jabr. Using twitter to follow trends beats the stock market. *NewScientist*, (2829), Sept. 2011. <http://www.newscientist.com/article/mg21128295.900-using-twitter-to-follow-trends-beats-the-stock-market.html>.
- [5] K. Levchenko, B. Meeder, M. Motoyama, S. Savage, and G. M. Voelker. Measuring online service availability using twitter. In *Proc. of the 3rd Workshop on Online Social Networks (WOSN 2010)*, 2010.
- [6] B. Liu. *Web Data Mining*. Springer, 2007.
- [7] Y. Matsuo, M. Okazaki, and T. Sakak. Earthquake shakes twitter users: real-time event detection by social sensors. In *WWW 2010: Proceedings of the 19th World Wide Web Conference*, 2010. <http://ymatsuo.com/papers/www2010.pdf>.
- [8] Y. Matsuo and M. Ishizuka. Keyword extraction from a single document using word co-occurrence statistical information. In *Proceedings of the 16th International FLAIRS Conference*, pages 293–296, 2003.
- [9] NIST/SEMATECH. 6.4.3.1. single exponential smoothing. e-Handbook of Statistical Methods. <http://www.itl.nist.gov/div898/handbook/>.
- [10] NIST/SEMATECH. 6.4.3.2. double exponential smoothing. e-Handbook of Statistical Methods. <http://www.itl.nist.gov/div898/handbook/>.
- [11] B. O’Connor, R. Balasubramanian, B. R. Routledge, and N. A. Smith. From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 122–129. AAAI Press, 2010.
- [12] M. F. Porter. An algorithm for suffix stripping. In K. Sparck Jones and P. Willett, editors, *Readings in information retrieval*, pages 313–316. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997. <http://tartarus.org/martin/PorterStemmer>.
- [13] E. R. Tufte. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, first edition, February 1997.
- [14] Twitter. #numbers, Mar. 2011. <http://blog.twitter.com/2011/03/numbers.html>.
- [15] K. Weil. NoSQL at twitter, 2010. Presented at Strange Loop conference.